

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

5

U.S. PATENT APPLICATION

FOR:

**ELECTRONIC MAIL METHOD AND SYSTEM USING
ASSOCIATED AUDIO AND VISUAL TECHNIQUES**

INVENTORS:

ADAM ROTH

GEOFFREY O'SULLIVAN

BARCLAY A. DUNN

ELECTRONIC MAIL METHOD AND SYSTEM USING ASSOCIATED AUDIO AND VISUAL TECHNIQUES

5

BACKGROUND OF THE INVENTION

Field of the Invention

The invention relates to electronic mail methods and systems and more particularly to electronic mail methods and systems using associated audio and visual techniques.

Background Information

Over the past decade, the use of electronic mail (email) has become a ubiquitous form of communication. Email has become among the primary methods of communications used in both business and social settings. Some of the advantages of communication by email are the ability to transmit messages instantly, the ability to transmit graphics, the ability to easily send messages to a large number of recipients, etc. Because of these (and other) advantages, transmitting messages by email is more efficient than other types of communication modes, such as telephone, facsimile, and the like.

Nevertheless, transmitting messages by email limits a sender's ability to convey emotions that may be associated with a message. For example, when sending a message, senders may wish to convey that they are angry, sad, happy, etc. Because the content transmitted by email is usually in a text format, it is difficult for a sender to convey emotions along with a message. Of course, certain punctuation (such as an exclamation point) or an emoticon may aid in conveying

a sender's emotion, but the value of using these symbols are limited. For example, it is difficult to use symbols to convey humor, impatience, urgency, empathy, etc.

U.S. Patent No. 6,064,383 to Skelley entitled "Method And System For Selecting An Emotional Prosody For A Graphical Character" is directed to a user interface element which enables a user to select an appearance for a character that corresponds with an emotion and an emotional intensity. Nevertheless, Skelley and the other existing prior art do not permit a full range of capability of expressing senses or emotions associated with a message transmitted over a network. For example, users are typically limited to a fixed number of characters, and are unable to easily display a likeness that resembles the user in a lifelike manner (with respect to audio and visual aspects) when sending such messages. Moreover, users cannot easily enter and choose among message content formats, be it text-based, audio, visual, etc., particularly with respect to the user's own likeness. Further, because transmitting messages by email has become so ubiquitous, it is desirable to have the ability to prepare such messages using a variety of devices and input techniques. Therefore, there is the need to devise a method and system in which a user can easily compose an email message with the ability to associate one or more multimedia components related to the message content. Additionally, there is the need to devise such a method and system wherein the user is not restricted to specific character choices, emotional displays, message formats or limited inputting techniques.

SUMMARY OF THE INVENTION

Accordingly, it is an object of the invention to provide a multi-mail system and method in which a sender may convey and a recipient can realize emotional aspects associated with substantive content of an email message by receiving a message that is more than textual in

nature. Voice recognition technology and programmatic relation of sound and graphics may be used to produce a talking image. In one embodiment, the image may include the user's own visual and/or audio likeness. In an alternate embodiment, the image may comprise any available visual and/or audio display selected by the user.

5 The multi-mail message may be inputted by a user in a text format and transposed into a format including the selected image and/or voice. In an alternate embodiment, a spoken message may be converted into a format including the selected image and/or voice. The formatted messages are then stored and/or transmitted via an email system or some other electronic network.

BRIEF DESCRIPTION OF THE DRAWINGS

10 The foregoing features and other aspects of the invention are explained in the following description taken in conjunction with the accompanying figures wherein:

15 FIG. 1A illustrates a network for supporting transmission of email messages according to one embodiment of the present invention;

 FIG. 1B illustrates one embodiment of the central controller used in the system shown in Fig. 1A;

 FIG. 2 illustrates the login process according to one embodiment of the invention;

 FIG. 3 illustrates the login process according to a second embodiment of the invention;

20 FIG. 4 illustrates the graphics and audio management process according to one embodiment of the invention;

 FIG. 5 illustrates the process of uploading an image or audio file according to one embodiment of the invention;

FIG. 6 illustrates the process of deleting an image or audio file according to one embodiment of the invention;

FIG. 7 illustrates the process of manipulating an image on a website according to one embodiment of the invention;

5 FIG. 8 illustrates the graphic add-on process according to one embodiment of the invention;

FIG. 9 illustrates the sound effect process according to one embodiment of the invention;

FIG. 10 illustrates the process of setting up or changing voice sample preferences according to one embodiment of the invention;

10 FIG. 11 illustrates the process of creating a new message according to one embodiment of the invention;

FIG. 12 illustrates the process of creating an audio portion of a message according to one embodiment of the invention;

15 FIG. 13 illustrates the image processing process according to one embodiment of the invention;

FIG. 14 illustrates the email creation process according to one embodiment of the invention;

FIG. 15 illustrates the text to audio synthesis process according to one embodiment of the invention;

20 FIG. 16 illustrates the process of receiving a spoken message by telephone according to one embodiment of the invention; and

FIG. 17 illustrates the sound quality processing process according to one embodiment of the invention.

It is to be understood that the drawings are for illustration only and are not limiting.

DETAILED DESCRIPTION OF THE INVENTION

5 System of the Invention

Fig. 1A shows one embodiment of a system incorporating the present invention. In this embodiment, the system includes a central controller 110, configured to receive information from one or more email senders 124 or to send such information to one or more recipients 134 at user interface devices 120, 130.

10 Central controller 110 preferably comprises a processor-based system that maintains databases and information relating to the preparation, transmission and receipt of multi-mail messages. A multi-mail message is an email message which has a multimedia component associated with the inputted message. The multimedia component may be in the form of audio content, image content (static or dynamic), graphic content or any other multimedia format.

15 Further, the multi-mail message may be transmitted with a textual component in addition to one or more multimedia components, or, in an alternate embodiment, the textual component is replaced by one or more multimedia components. Although the multimedia component may be added to a message to facilitate conveying a user's emotion associated with a message, the component(s) may be utilized to simply more effectively transmit a sender's message.

20 As described more fully below, a multi-mail message may be prepared in a number of different ways. In one embodiment, an email message may be prepared by typing in a textual message at a sender's user interface device 120 that is capable of receiving such text, and associating an accompanying audio and/or video component. In another embodiment, an email

message may be composed by telephone, or some other device that is capable of receiving an audio input. In one embodiment, user interface devices 120, 130 may be a computer comprising one or more central processing units (CPU), one or more data storage devices, a monitor, a keyboard and/or any other components that may allow a user to implement the commands of the software and hardware functions described herein. Further as Fig. 1A illustrates, a sender user interface device may be a personal digital assistant (PDA) 120-1, telephone 120-2, or any type of wireless or land line device that is capable of transmitting textual, graphical or audio messages, such as facsimile machines, beepers, etc. Similarly, any devices that are capable of communicating an audio and/or video message (with or without transmitting textual data) are suitable for displaying email messages to a receiving party.

Referring to Fig. 1B, central controller 110 stores information received from senders or recipients in user/account database 352. Further, information concerning animations, images and audio components which may be associated to a specific message and/or a specific user is stored in databases 354, 356, 358. As described more fully below this information is used to compose, send and receive multi-mail messages which convey the sender's message in a manner that is not only textual. In some cases, the multi-mail message may have no textual component at all. The data stored in databases 354, 356, 358 effectuate the transmission of such message formats. The structure of certain embodiments of the central controller 110 is described below in connection with Fig. 1B.

Senders 124 include individual(s) wishing to compose and/or send a multi-mail message to recipient(s) 134. As described above, a multi-mail message is an electronic message that has at least one multimedia component associated with the message. In one embodiment, users (senders 124 and recipients 134) register, through user interface devices 120, 130, with the

service which performs the email processes as further described below. In one embodiment, the user interface device may be the user's computer or internet access device. In another embodiment, the user interface device may be a telephone, PDA, facsimile machine, etc.

Senders 124 can input email message information in various ways. For instance, users may send such messages electronically by means of the internet, world wide web, intranet, or some other network. This may be accomplished by inputting a message through sender device 120 which is in communication with the network. In an alternate embodiment, the sender device may be in communication with some other type of network, such as the public switched telephone network (PSTN). Two ways in which a sender may input such information via the PSTN for receipt by central controller 110 include (1) telephone answering services at central controller 110 that provide audio recording capabilities and, in some instances, that permit interaction between the user and the system based on touch tone or voice input provided by the user; or (2) telephoning live operators at central controller 110, to verbally provide information which is entered into the system via operator terminals.

Figure 1B illustrates one embodiment of the central controller 110 for a system according to the present invention. As shown in Fig. 1B, central controller 110 includes central processing unit (CPU) 240, random access memory (RAM) 220, read-only memory (ROM) 230, and large capacity storage device 250. CPU 240, preferably comprising a conventional microprocessor such as an Intel Pentium Processor, is electronically coupled to each of the central controller's other elements.

CPU 240 executes program code stored in one or more of RAM 220, ROM 230 and storage device 250 to carry out the functions and acts described in connection with central controller 110. CPU 240 preferably comprises at least one high-speed digital data processor

adequate to execute program modules for executing email transmissions by others 124, 134.

These modules are described in connection with Figs. 2-17. CPU 240 interacts with RAM 220, ROM 230 and storage device 250 to execute stored program code according to conventional data processing techniques.

5 User interface devices 120 and 130 comprise devices for allowing central controller 110 to communicate with multi-mail senders 124 and recipients 134. Communication between these user interface devices 120, 130 and the controller 110 is preferably electronic by means of the internet and preferably includes a conventional high speed modem employing known communication protocols capable of decrypting encrypted data received from the interface user devices 120, 130.

10 Large capacity storage device 250 contains transaction processor 260, user/account database 352, animation database 354, image database 356 and audio database 358. Transaction processor 260 maintains, determines and accesses data stored in databases 352, 354, 356, 358, and handles multi-mail message information for transmission between senders 124 and recipients 134 as described in connection with Figs. 2-17. Transaction processor 260 may comprise a separate, conventional CPU/microprocessor, or a portion of the operating function of CPU 240.

15 User/account database 352 contains data about senders and recipients that register with the multi-mail service described herein. The data is used to store subscriber information including name and email address(es), as well as information for accessing audio, video and animation files, where appropriate. Animation database 354 contains information relating to animations associated with a sender or sender's message. Animation information includes data relating to a dynamic graphic associated with a sender's message. Image database 356 contains

information relating to images associated with a sender or sender's message as well as data for controlling the audio and video aspects of a dynamic image. Image information may include any dynamic or static graphic associated with a sender's message. Audio database 358 contains information relating to audio data associated with a sender or sender's message. In one
5 embodiment, the image may be a graphic of a user's, or someone else's, face. In an alternate embodiment, an image may be any other type of graphic. Audio information may include any voice(s) or sound(s) associated with a sender's message. As described more fully below, the animations, graphics and/or audio components associated with a sender's email message may include the likeness of the sender. In another embodiment, these multimedia components may
10 have little or no resemblance to the sender. The process of using data from user/account information database 352, animation database 354, image database 356 and audio database 358 to prepare, send and receive messages, as well as the other operations of the system described with reference to Figs. 1A and 1B, is represented in the flowcharts of Figs. 2-17, described in detail below. Central controller 110 uses user/account information database 352, animation
15 database 354, image database 356 and audio database 358 to execute multi-mail preparation and transmission processes.

Login Process

The login process is the preliminary process in which users (senders 124, recipients 134)
20 provide information to effectuate access to the multi-mail system and serves at least the following purposes: (1) storing address information to enable proper routing of a multi-mail message; (2) transmitting messages (electronic, telephone, mail, etc.) by the multi-mail server to users such as new messages, return receipts, etc.; (3) effectuating login by comparing identifier

information entered by users during the login process to the stored personal identifier information provided at registration; and (4) enabling users to select a link so that message information and/or user information may be stored.

Referring to Fig. 2, the login process begins at step 1000 where a user accesses the multi-mail program stored in central server 110 via user interface 120, 130. In one embodiment, this is accomplished by typing a universal resource locator (URL) associated with the multi-mail site. A determination is then made, in step 1100, as to whether the user is an existing multi-mail user or a first time visitor. If the user has previously registered with the multi-mail site the user logs in by entering a username and associated password. This information is compared to the data previously stored in user/account database 352. Once the information is verified, the user may further access the multi-mail site.

Once the login information is verified as being correct, the user may access several options including, but not limited to, managing graphics and audio files (step 1111), creating a new multi-mail (step 1112), accessing multi-mail via a free multi-mail account (step 1113) and creating or editing user or multi-mail preferences (step 1114). Various aspects concerning managing graphics and audio files (step 1111) are described below with reference to Figs. 4-11. Various aspects concerning creating a new multi-mail message (step 1112) are described below with reference to Figs. 12-17. Accessing multi-mail via a free email account (step 1113) allows a user to retrieve email sent to a personal email account at gabmail.tv, as well as email sent to them at other POP accounts. Finally, creating or editing preferences (step 1114) allows a user to select default settings associated with the user's account and/or messages. Such settings may include default audio, image and animation attributes, multi-mail routing information (i.e., priority information, forwarding information, etc.) and the like. The default settings may apply

to all or only a sub-group (such as those to a specific recipient 134) of multi-mail messages created by a user.

If the user has not previously registered with the multi-mail site, a new username and password is established by the user and is stored in user/account database 352 (step 1120). The user may then create preferences associated with the user's new account and/or individual messages (step 1114). Upon creating these preferences, the user may manage graphics and audio files (step 1111), create a new multi-mail (step 1112), access multi-mail via a free multi-mail account (step 1113) and/or edit user or multi-mail preferences (step 1114).

Telephone Access Process

In addition to accessing the multi-mail system by computer, a user may access the system by telephone. (It should be noted that access to the multi-mail site is not limited only to computer and telephone access. These are merely representative samples of devices that may access the system by either some graphical user interface (i.e., computer) or by audio interface (i.e., telephone).

In step 2000, a user accesses the multi-mail system through the PSTN by dialing a telephone, such as a toll free 800 number. Upon accessing the system, the user specifies, in step 2100, whether they have an existing account. If an account does not already exist, the user may make such an indication by pressing a key on the telephone keypad (i.e., "*") or by speaking a predetermined word as requested by a multi-mail interactive voice response unit (step 2120).

If, however, CPU 240 of central controller 110 determines that a user account already exists, the user's identification code and password are entered and compared to the data stored in user/account database 352. Upon verification that the user's login information is correct, the

user may access new (or old) received multi-mail messages (step 2111), compose a new multi-mail outgoing message (step 2112) or change the user's personal settings (step 2113). Changing a personal setting may include recording a new audio component to be associated with multi-mail messages (step 2113.1). This new audio component may be stored in audio database 358 in place of or in addition to any audio components that the user already has stored. Additional personal settings, such as choosing a default image (step 2113.2), listening to other voice samples (2113.3), or choosing background sounds (2113.4), may be activated by the telephone system. Other options, such as recording samples of a laugh, a whisper, a shout, or the like are available.

Graphics and Audio Management Process

Referring to Fig. 4, various audio and graphics capabilities of the multi-mail system are illustrated at step 1111. For example, at step 1111.1, image or audio files may be uploaded. As described above such files are stored in and may be accessed from any one or more of animation database 354, image database 356 and audio database 358. These files may be .wav, .jpeg, .mpeg, or any other type of file containing audio or video data. In addition to storing such files, these files may be deleted from database 354, 356, 358 at step 1111.2.

Images that can be accessed by a user may also be manipulated by the user. In one embodiment, the manipulation may take place at the multi-mail website (i.e., central controller 110) and may include enlarging or shrinking an image, reorienting the image, changing the colors of the image, tiling or cascading the image, etc. (see Fig. 7 and related description below for more detail). Further, at step 111.4, the graphics manipulation enables graphic add-ons which are described more fully below with reference to Fig. 8.

At step 1111.5, sound effects may be stored in and/or accessed from audio database 358.

Sound effects may include changing the pitch, rate of speech, etc. of an existing message.

Further, sound effects may include choosing from a library of common sounds including, but not limited to, the sound of breaking glass, wind blowing, door slamming, dog barking, etc. Another

5 library of sound effects may be the voices of celebrities or other famous people. In other words, the sound effects may result from changes to a user's voice, another person's voice or any sound related or unrelated to a person's voice. Finally, at step 1111.6, a voice sample preference may be created or changed. Such preferences (1111.6) as well as the other graphic/audio management steps (1111.1-1111.5) are described more fully below with reference to Figs. 5-10.

10 **Uploading an Image or Audio File**

An image or audio component may be associated with a multi-mail message by accessing from database 354, 356, 358 the file(s) which store such data. Referring to Fig. 5, a user who wishes to upload an image or audio file (step 1111.1) may, in one embodiment, browse to or type
15 in the file location on a local computer (or some other user interface device 120 which may browse or access such files) to locate such data files (step 1111.11) or terminate the upload session. Once the appropriate file is located, the user may select and upload the data of such file (step 1111.12). The sound, image or graphic component may then be previewed at step 1111.13. If the component is not approved (step 1111.14), the user may return to the browse mode (step
20 1111.11). If, however, at step 1111.14, the component is approved, it is associated with the message and the user returns to component management menu (step 1111). Further, a keyword may be associated with the file so that the contents of the file may be accessed via telephone 120-2.

Deleting an Image or Audio File

Referring to Fig. 6, a user may choose to delete one or more audio, image or animation files (step 1111.2). This is accomplished by first selecting an audio, image or animation one or more available files stored in database 354, 356, 358 associated with the user's account (step 1111.21). Once the file(s) are selected, the user is asked to confirm the deletion of the file(s) (step 1111.22). If the selected file(s) are not confirmed for deletion, the user may choose to select other file(s) for deletion (or terminate the file deletion process). If, however, at step 1111.22, the user confirms deleting the selected file(s), those file(s) are deleted from one or more of databases 354, 356, 358 (1111.23) and the user receives confirmation of such deletion (1111.24). In one embodiment, the user is then returned to the component management menu (step 1111).

Manipulating an Image

As described above, a user may wish to manipulate an image for association with a multi-mail message. Such manipulation begins at step 1111.3 of Fig. 7. Image manipulation enables a user to easily change a stored image. Although a user may change a stored image in countless ways, some of the manipulation options are as follows: (1) include stock or own animation which could, for example, include a mini-movie in the background behind the talking head (step 1111.31); (2) drag an image of a face into a goofy shape, (step 1111.32); (3) enlarge eyes or exaggerate other image features (step 1111.33); (4) provide color enhancement to an image or video (i.e., technicolor, grainy black and white, sepia, etc.) (step 1111.34); (5) enable image

movement, such as winking (step 1111.35); (6) change component background (i.e., street scenes, circus, ships, psychedelic, subtle, stripes, antique, etc.) (step 1111.36); (7) drag other objects into image or video (step 1111.37); and (8) move or alter body parts (step 1111.38). The manipulations listed above are merely representative examples of the effects that may be applied to a multi-mail image or video components and are not meant to be an exclusive list of manipulations available. After completing the image/video manipulation (step 1111.3), the user may return to the component management menu (step 1111).

Graphic Add-On Process

As shown in Fig. 8, a user may also wish to apply graphic add-ons to a multi-mail message (step 1111.4). These add-ons enable a user to introduce additional images into a graphic component. The additional image may already be stored in image database 356, may be accessed by some other database storing such images or may be scanned into the user's local user interface device 120. Although a user may access countless graphic add-ons, some of the add-ons that may be applied to a multi-mail message are as follows: (1) applying various body types, including different sizes and proportions, color (skin, hair, etc.), features, etc. (step 1111.41); (2) including different types of clothing and accessories to images of persons (step 1111.42); (3) applying different tints to the graphic (step 1111.43); and (4) applying different surroundings and wallpapers to the graphic (step 1111.44). The graphic add-ons listed above are merely representative examples of the add-ons that may be applied to a multi-mail image or video component and are not meant to be an exclusive list of the available graphic add-ons. After completing the graphic add-ons process (step 1111.4), the user may return to the component management menu (step 1111).

Sound Effect Process

As shown in Fig. 9, a user may also wish to apply sound effects to a multi-mail message

(step 1111.4). These sound effects enable a user to include either speech or non-speech sound

that is associated with a multi-mail message. In one embodiment, the sound effect may simply

be the user stating the textual contents of the multi-mail message using his or her normal voice.

In another embodiment, the sound effect may be the user's speech attenuated to convey an

emotion associated with the message. In yet another embodiment, the sound may be the speech

of a person other than the user. Finally, in another embodiment, the sound effect may be some

noise that is not speech. Some examples of the sound effects that may be applied to a multi-mail

message are as follows: (1) applying joke or gag sounds (step 1111.51); (2) including

background music (step 1111.52); and (3) choosing from generic voice samples (step 1111.53).

The sound effects listed above are merely representative examples of the sounds that may be

applied to a multi-mail image or video component. The sound effects may already be stored in

audio database 358, may be accessed by some other database storing such sounds or may be

recorded into local storage of a local user interface device 120 that is capable of receiving sound

data. After completing the sound effects process (step 1111.5), the user may return to the

component management menu (step 1111).

Establishing or Changing Voice Sample Preferences

Referring to Fig. 10, voice sample preferences may be established or changed (step

1111.6). A voice sample is any voice that a user may store and/or access for association with a

multi-mail message. As described more fully above, the voice may be that of a user, a variation of a user's voice, or someone else's voice.

In one embodiment, a voice sample may be stored in audio database 358 by telephone 120-2 or any other device having an audio input and output in communication with central controller 110. At step 1111.61, a user initiates the voice sample call in process by accessing a telephone number (such as a toll free telephone number) which is associated with the multi-mail message service. The user logs into the system by providing a username and password (step 1111.611). The user then navigates to a voice sample recording section of the system (step 1111.612). In one embodiment, this is accomplished by pressing the appropriate button on a keypad in response to the system requests. In an alternate embodiment, an interactive voice response unit may be used. At step 1111.613, the voice sample is recorded. Recording devices may be part of the telephone system. The user will read a statement into the phone that will provide a broad enough spectrum of sounds to allow the automated synthesis of any text the user may type in the future. The user may then review the recorded sample at step 1111.614. If the user does not approve of the recorded sample (step 1111.615), the process returns to step 1111.613 where the voice sample may be re-recorded (or the process may be terminated). If, however, the recording is approved, the user may confirm such approval (step 1111.616). The system saves and then parses the voice sample into phonemes (step 1111.617) and organizes, sorts and stores the phonemes (step 1111.618). More specifically, the system will (1) analyze sample for breaks (silence) and divide it into words at breaks; (2) divide words into syllables; (3) divide syllables into component vowel and consonant sounds, or phonemes; (4) identify each in the set of phonemes thus derived from voice sample by comparing to a standard set of spoken phonemes; and (5) eliminate duplicates, enumerate any missing phonemes, and organize

collection for future use. At this point, a person's voice is stored in such manner that the multi-mail system may transmit a voice-based message to a recipient 134. In another embodiment, an audio component may be associated with a textual message by converting the textual message to an audio one. This may be accomplished by utilizing a text to speech converter which is known in the art.

In addition to calling in a voice sample (step 1111.61), a user may choose a voice sample from a generic voice sample stored in, for example, audio database 358 (step 1111.62). Further, an audio file may be uploaded as described above (step 1111.1).

10 **Creating a New Message**

Referring to Fig. 11, a user may create a new multi-mail message at step 1112. This is accomplished by (1) creating an audio portion (step 1112.1), (2) selecting a preexisting image or uploading a new image (step 1112.2), (3) processing the image (step 1112.3), and (4) synthesizing the multi-mail message (step 1112.4). Steps 1112.1-1112.4 are described in greater detail below with reference to Figs. 12-17. It should be noted that although the multi-mail system supports all types of multimedia components (i.e., audio, image, animation, etc.), there is no requirement that more than one component be selected by the user. For example, the user may wish to add only an audio component, in the form of the user's voice, to an otherwise standard text email. Alternatively, the user may want to show a picture of an object in association with the message.

Once a multi-mail message has been composed with the appropriate associated audio, image and/or animation components, the user may preview the multi-mail message at step 1112.5. If the user does not approve the message (step 1112.6), the multi-mail message

composition process may be terminated or the user may return to step 1112.1 where the message may be edited or created anew in part or in toto. If, however, at step 1112.6, the user approves the multi-mail message, the user may choose to transmit the multi-mail message (step 1112.7). In one embodiment, the message is transmitted as soon as the multi-mail message is approved.

5 In an alternate embodiment, the user may choose an option to have the message sent either automatically or manually at some future time. In any event, when the message is transmitted, the user may select an option to receive a confirmation (step 1112.8) when the transmitted message has been sent by the sender 124 and/or received by the recipient 134.

10 **Creating an Audio Portion of Message**

As described above, a multi-mail message may include an audio component to more effectively convey a message. Creating the audio component of a message begins at step 1112.1 of Fig.12. Entering the audio component of a multi-mail message into the mail system may be accomplished in at least two ways. In one embodiment, a sender 124 may enter text via user interface device 120 (step 1112.11). CPU 240 of central controller 110 may then synthesize the typed text into audio data which is stored in audio database 358 (step 1112.111). This is accomplished by using a text to speech converter in conjunction with the parsed voice sample stored in audio database 358 that is selected by sender 124.

15 In an alternate embodiment, the sender 124 may submit a spoken message by telephone 20 120-2 or some other device which has an audio input and output capability and is in communication with central controller 110 (step 1112.12). Next, central controller translates the sound to text by accessing a speech to text converter. The conversion is performed and saved at

step 1112.121. Sound quality processing is performed at step 1112.122 and the audio component creations process is complete (step 1112.1).

Image Processing

As described above, a multi-mail message may include an image component to more effectively convey a message. Creating the image component of a message begins at step 1112.3 of Fig.13. The image processing module enables a user to store his or her own likeness or that of another (using a digital camera, developing a standard photograph digitally onto a CD, scanning an image, or using some other digital input device) to more effectively convey a message.

Further, the module enables the user to combine the audio component with a dynamic image component, to enable a recipient 134 to receive a message wherein a person's image and voice is conveying a message.

At step 1112.41, the user indicates to CPU 240 of central controller 110 whether the image is a face (step 1112.41). If the image is a face, the eyes and mouth are located (at step 1112.411). An image recognition application that looks for almond shapes (eyes), oval diamonds (mouths), or ovals (faces) may be either custom written or existent and known to those skilled in the art. Further, mouth shape frames are generated for each corresponding message vowel and consonant (step 1112.411) and eye blink frames are generated as well (step 1112.4112). The system includes custom written code that utilizes an algorithm or set of algorithms to make minute changes to an image, save in a new file, make more minute changes, save again in a new file, and so on until the collection of images together form the frames of animation to look like a speaking face. The images would then be compiled in order forming an animated sequence. If the image is not a face (or even if it is) the user may use a custom-written

web-based application to denote regions of their image they would like to animate, and then select from a pre-determined collection of types of animation (blinking eyes, moving mouth, nodding head, or limited set of other non-facial movements) to apply to those regions of the image.

5

E-Mail Creation Process

Referring to Fig. 14, the multi-mail message is created at step 1112.4. In one embodiment, the text is first entered into the multi-mail message file (step 1112.41). Next, an associated sound clip is parsed into phonemes (step 1112.42) if the audio component was previously entered by telephone or a similar audio device. If, however, the message was originally typed in (i.e., textual), the phoneme information is recalled from the parsed typed text which is stored in image and audio databases 354, 358 (step 1112.41a).

Next, at step 1112.43, a phoneme/time sequence is composed by CPU 240. Further, a matching mouth frame/time sequence (step 1112.44) and a non-matching eye blink frame sequence (step 1112.45) are composed. The sequence of image frames is composed to match the sequence of phonemes, in order and in individual frame duration. The image sequence and sound information are then combined into a synchronized clip which is stored in animation database (step 1112.46). Finally, the animation clip is embedded into hypertext markup language (HTML) in a multi-mail message for transmission to a recipient's user interface device

130.

Text to Audio Synthesis Process

Referring to Fig. 15, the text to audio synthesis process is described. A sender most frequently enters an email message by typing in a text message which is displayed as a graphical user interface on the sender's user interface device 120 and then sent to a recipient's 134 user interface device 130. Because multi-mail messages have a multimedia component associated with the text message, the central controller converts the text message to an audio/image display to accompany and/or replace the text display. The synthesis enables such display to look and sound life-like. The text to audio synthesis begins at step 1112.111. First, central controller 110 parses the textual message into sentences. This may be accomplished by recognizing punctuation provided by the user or some other algorithm known in the art. Next, at step 1112.112, voice modulation controls are assigned to sentence end. This enables appropriate inflection and pausing to be accomplished as sentences end.

At step 1112.1113, the text is then parsed for voice modulation controls, or custom codes to indicate a variety of voice inflections or adjustments to be made on that sound. The sentence type (i.e., question, compound sentence, etc.), as well as emphasis commands provided by the sender 124, may dictate how the text is parsed. Next, phoneme and modulation information is sequenced and translated into a sound sequence (steps 1112.1114 and 1112.1115) and the voice is modulated pursuant to the modulation controls or commands (step 1112.1116). Using sound effect plug-in tools known to those skilled in the art, such as Mark of the Unicorn's Spectral Effects, the voice can be modulated in terms of pitch (higher or lower), formants (phoneme energy), or speed (time expansion or compression), to indicate a range and intensity of emotion. Finally, the sound quality is processed, at step 1112.21, and the associated multi-mail message audio file is compiled (step 1112.1117). The user may then send the message with the associated

audio file; the associated text may be optionally transmitted. Further, as described above, an associated image file may also be associated with the message and transmitted.

Receiving Spoken Message by Telephone

5 Referring to Fig. 16, the process of receiving an audio message via telephone 120-2 or some other device having audio input and output capabilities is described. At step 1112.12, a sender 124 submits a spoken message by telephone 120-2. In one embodiment, the voice sample may be stored in audio database 358.

10 The user first accesses the system by calling a telephone number (such as a toll free telephone number) which is associated with the multi-mail message service. The user logs into the system by providing a username and password (step 1112.121). The user then navigates to a voice sample recording section of the system (step 1112.112). In one embodiment, this is accomplished by pressing the appropriate button on a keypad in response to the system requests. In an alternate embodiment, an interactive voice response unit may be used. At step 1112.123, 15 the voice sample is recorded. The multi-mail message service system contains software whose function it is to digitize (i.e., record) sound. At the appropriate signal from the multi-mail message service, this software is invoked to record the user's voice sample. The user may then review the recorded sample at step 1112.124. If the user does not approve of the recorded sample (step 1112.6), the process returns to step 1111.123 where the voice sample may be re- 20 recorded (or the process may be terminated). If, however, the recording is approved, the user may confirm such approval (step 1111.126). The system then performs the necessary recorded message processing (step 1112.127) as described in detail above with reference to steps 1111.617 and 1111.618 of Fig. 10.

Sound Quality Processing

Fig. 17 illustrates the sound quality processing of the system shown in Figs. 1A and 1B.

Sound quality processing is comprised of the different modules that central controller 110

5 performs to the audio component of a multi-mail message to ensure that the sound has realistic

qualities. At step 1112.211, the equalization process is performed to adjust individual sound

wave frequencies. Next, at step 1112.212, central controller 110 performs the compression

process, smoothing out sound wave peaks and compressing audio frequencies. Next, at step

1112.213, noise gating is performed, which further improves the sound quality by only

10 permitting the audio component to pass if it reaches a certain user defined threshold, and an

audio limiting process is performed by central controller 110, which limits the audio signal at a

user defined level. These audio processes, eliminating background noise and other distortions

and improving overall audio quality using methods known to those skill in the art, provide a

high quality audio component to be communicated to a recipient 134.

15 It will be apparent to those skilled in the art that various modifications and variations can

be made in the system and processes of the present invention without departing from the spirit or

scope of the invention. Thus, it is intended that the present invention cover the modifications

and variations of this invention provided they come within the scope of the appended claims and

their equivalents. In this context, equivalents means each and every implementation for carrying

20 out the functions recited in the claims, even if not explicitly described herein.